DOCUMENT RESUME

ED 163 404 CS 004 503

AUTHOR Lynch, Michael
TITLE A Remedial Programme Based on a Diagnostic Test for Irish Pupils.
PUB DATE Oct 77
NOTE 11p.; Paper presented at the Annual Meeting of the Reading Association of Ireland (2nd, Dublin, Ireland, October 6-8, 1977)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Diagnostic Teaching; *Diagnostic Tests; Elementary Secondary Education; Item Analysis; *Profile Evaluation; *Reading Diagnosis; *Remedial Reading Programs; Testing Problems; *Test Validity

ABSTRACT

The way in which diagnostic reading test scores were used in planning a remedial program is examined in this report. Among the topics discussed are the administration of a diagnostic test in 13 Irish schools; the organization of a remedial program for students identified as having severe reading problems; weaknesses of diagnostic tests currently in use; the concept of profiling and cautions to be considered in developing profiles; and the use of Rasch-Wright procedures, which were used in extensive analysis and calibration of the items on the diagnostic test. To show the types of analyses made and the types of information incorporated into the remedial program, an individual student's scores on the diagnostic test are reported for such categories as knowledge of homophones and prefixes. It is noted that students' involvement in focusing on individual items provided a sense of purpose and contributed to the success of the remedial program. (GW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED163404

A REMEDIAL PROGRAMME BASED ON A DIAGNOSTIC TEST FOR IRISH PUPILS

Michael Lynch

Education Department, University College, Dublin 4.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Michael Lynch

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

INTRODUCTION

In 1976 primary schools in a provincial Irish centre cooperated with the writer in testing pupils in standards V and VI with a view to develop a diagnostic reading test which would be closely related to the content of instruction. Thirty teachers in thirteen schools participated and just over one thousand students were tested - the total school population in the local schools at this level. Written responses were obtained for all students on seven hundred items covering many aspects of vocabulary. The response matrix of almost three-quarters of a million entries was extensively item analysed by means of Rasch-Wright procedures.

The purpose of this paper is not to report on the construction of the diagnostic test but to discuss a remedial programme with twenty students in one of the schools who were identified as having severe reading problems. The remedial programme began in December 1976 and continued until June of 1977 with the assistance of the Irish Department of Education. Pupils were withdrawn from their normal class for one and a half hours three mornings per week in groups of ten. They also received individual instruction at other times during the day, apart from normal class instruction. Every effort was made to utilize the test information diagnostically in planning the remedial programme.

It is hoped to discuss problems inherent in this approach in general terms and to illustrate ways in which information from the test was used to plan instruction by reference to a student's responses.

CS004503

Global cognitive measures such as the Stanford-Binet and the Wechsler Intelligence Scale for Children provide very limited information with which to guide subsequent instructional programmes. Yet clearly there is a demand, particularly from the field of special education, for diagnostic profiles which will reveal intraindividual differences and indicate areas in need of remediation (Ward, 1970). Unfortunately instruments constructed with this end continually in mind have had only mixed success. The Illinois Test of Psycholinguistic Abilities (Kirk et al., 1968) seeks to analyze the child's psycholinguistic abilities along three dimensions deemed to be important: (a) channels of communication (auditory-vocal; visual-motor), (b) processes (reception; organisation; expression), and (c) levels of organisation or complexity (automatic level; representational level). The classification of specific functions should give clues to possible remedial action (Paraskevopoulos & Kirk, 1969). In her study of young fluent readers Clark (1976) observed that the results of the ITPA added little of diagnostic value for the fluent readers when taken in association with their scores which were already known from other language tests, and when compared profiles of children known to have reading difficulties. She urges caution in planning remedial programmes based on deficits shown on such tests. A comparison of the results for the fluent readers at six years and of the backward readers of average intelligence of nine years was striking not for the differences in the profiles but for their similarity !

The Frostig Developmental Test of Visual Perception claims to measure five relatively unrelated but important areas of visual perception (Frostig, 1963). The test has been vigorously attacked and the claims that identification of strengths and weaknesses followed by appropriate training exercises resulted in improved reading scores have been greeted with scepticism (Pumfrey, 1976). Tew (1976) has stated that over ten years it has come to be regarded uncritically as an efficient diagnostic instrument for the identification of perceptual handicaps which are presumed to contribute to difficulties in acquiring skills. Tew's own study in South Wales is cited as an example of research which does not substantiate the Frostig claims.

Olsen (1968) and Smith & Marx (1972) in the United States found no relationship between Frostig subtest scores or the perceptual quotient and the reading ability of normal school children. Tew cites the Scandanavian study of Neilsen & Ringe (1969) where it was observed that the Frostig test did not discriminate between retarded and normal readers. In fairness it must be stated that subsequent issues of Remedial Education carried correspondence critical of Tew (e.g., Topping, 1976).

Profiles

Since profiling was central to the diagnostic programme under consideration, some observations on the concept would seem apposite. The notion of profiling across different aspects of reading performance carries enormous intuitive appeal. Zig-zag lines linking high and low points on a graph do produce a most striking graphical display. Doe in a recent article in the Times Educational Supplement (Doe, 1977) reproduced one such appealing profile for the Richmond Tests of Basic Skills (Hieronymus et al., 1974). This test has been adopted by Cumbria and a special grid for Richmond test scores included in the county's new record card. One hopes that the teachers will treat such profiling with professional caution. The special problems relating to differential diagnosis have always been clear to test constructors but unfortunately in popular treatment they have been either ignored or glossed over. In the zig-zag treatment the standard error attached to the different scores tends to be ignored. Bookbinder (1976) says "most teachers are about as interested in the standard error of a test as most politicians are in the standard error of a vote-predicting public opinion poll (p.21)."

A more sophisticated version of the zig-zag treatment regards attainment on a subtest not as a fixed point but as a band varying within one or two standard errors of a midpoint. This enables statements to be

made with precisely stated statistical confidence about an individual's score. A profile across ssubtests according to this scheme will appear as a series of vertical bands arranged across the page. Where the bands overlap little significance can be attached to score differences.

Visual discrepancies between high and low points do not necessarily indicate obvious areas of weakness. One has to consider the inter-correlations between the subtests as well as subtest reliabilities (Anastasi, 1968). Unfortunately unwarrented interpretations are far too common: Savage(1968) says "interpreting subtest scores without regard for statistical considerations is a common practice which should be strongly discouraged, if not banned (p.35)".
Even greater difficulties arise when different norm groups are utilized in the construction of the subtests and when profiling is then attempted across the composite tests. The APA Standards (1966) urge great caution with such mixed scales. Reputable tests will warn of the dangers and the mistakes of attaching too much weight to every discrepancy in scores e.g. the manuals of the Edinburgh Reading Tests. Less acceptable tests may not even report the intercorrelations and separate subtest reliabilities and seek to ignore the problem! AERA's Technical Recommendations state that 'if the manual recommends that diagnostic interpretation be based on the test profile, evidence should be provided that the profile patterns used are valid discriminators (p.25, C.11).'

Reporting in terms of spelling or reading ages is even more suspect (Vincent, 1974), despite their popularity with teachers. The objections are fully discussed in Vincent & Cresswell (1976). Obviously the preferred units of teachers do not coincide with those of the research workers (Bookbinder, 1976) !
Problems in score interpretation are not confined to measures of attainment. In the field of guidance and counselling some interest questionnaires provide ipsative scores - multiple scores summing to a constant. These scores are open to even greater misinterpretation (Closs, 1976).

Rasch Scales

Test items were extensively analysed and calibrated by Rasch-Wright procedures. A value was assigned to each of the 700 items - its 'difficulty' on an interval scale extending from -6 to +6.
The Rasch model is probabilistic in nature and lies within the context of latent trait theories of which it is a special case. In latent trait terminology person attainment is described as person 'ability'. When person ability is exactly the same as item difficulty the person has got a 50/50 chance of correctly answering an item.
Both item difficulty and person ability values are on the same interval scale. The chances of success can be precisely stated.

From person performance on a given cluster of items, a subtest, it is possible to estimate person ability values corresponding to different score groups, just as one can convert raw scores to standard scores in a conventional test manual. In the case of Rasch analysis the ability values are not standard scores and grade norms but values which lie along the same interval scale as the item difficulties.
A good introduction to this method of scaling is provided in the report to the Schools' Council by Willmott and Fowles (1974); for a very recent account within the context of advances in item analysis consult Baker (1977). The method has been applied to test construction in the U.S.A. in the Woodcock Reading Mastery Tests (Woodcock, 1973).
Rasch would claim that the result of this kind of item calibration is a measure of person ability which is independent of the items attempted. This claim has important consequences for instructional programmes. Items presented to a person after the treatment period need not be the same as before, they can even be at a higher level of difficulty. It is only necessary that they fit the Rasch model.
This claim was stringently tested in the preparation of the main dissertation and certainly was quite convincing for the data.

Application

Consideration of an individual student may serve to relate the foregoing discussion to the remedial programme.
One student was 13½ years at the beginning of the remedial programme. His profile was based on his results in the district-wide testing

programme when he was in sixth standard. At the end of that year he was independently assessed by the local technical school and was considered too low in reading attainment to benefit from the courses. He repeated sixth standard in the primary school and it was during that year that he took part in the remedial classes.

His ability (attainment) on a subtest of Homophones was -0.06 on the scale. The average Rasch value for this subtest was +0.81 for V standard (N = 494), and +1.39 for VI standard (N = 520). His value was thus well below the mean for sixth class. Let us consider in some depth the meaning of the score of -0.06. In the Homophones subtest the student was asked to write 'a word which sounded exactly the same as the given word, but had a different meaning and spelling'. Sufficient examples were provided, e.g. flea, flee; cent, sent; hymn, him. The subtest had two parallel forms with a subtest reliability of +.92. Combined raw scores on the two parallel forms correlated +.79 (N = 541) with combined raw scores on the Vocabulary and Spelling subtests of the Drumcondra English Test, Level III.

At the level of individual items the student's ability value of -0.06 would give him a probability of .50 of responding to the following items:

	Item	Correct Response	Item Difficulty	Student's Response
1.	nun	none	-0.027	incorrect
2.	scene	seen	-0.133	correct
3.	chord	cord	-0.113	correct
4.	haul	hall	-0.092	correct
5.	prey	pray	-2.064	correct
6.	cruise	crews	+3.019	incorrect

For the first four the student had a probability of about .5; for no.5 about .90 and for no.6 'cruise' only about .05 because of the item's difficulty.

In the subtest Prefixes the student was required to encircle the prefixes for 27 items after sufficient examples and explanation. The student's Rasch value on this subtest was +.60. Bear this in mind in considering the items below:

	Item	Prefix	Difficulty	Student's Response
1.	overgrown	over-	-2.723	correct
2.	enmesh	en-	-1.584	"
3.	cooperate	co-	+0.711	incorrect
4.	circumnavigate	circum-	+0.814	"
5.	perinatal	peri-	+1.556	"
6.	contradictory	contra-	+2.606	correct
7.	infrastructure	infra-	+2.176	incorrect

At the beginning of the programme the student showed the following deficits on well known tests:

Neale Analysis of Reading Ability:	minus 5yrs. 5mths.
Schonell Spelling Form A	minus 5yrs.
Marino	minus 4yrs 9 months.

Similar analyses were made for all students in the remedial class and the information incorporated in the remedial programme. Their scores and responses were available at all times to the students and they were encouraged to examine their responses and to relate them to the item difficulties.

Analysis of performance took into account responses to individual items rather than summed scores. Clinical observations were made on persistence and individual styles of response as has been advocated by Eysenck (1967). The students' involvement provided a sense of purpose and contributed to the success of the remedial programme. Vygotsky (1962) believed that the child has little sense of purpose in writing and has only a vague idea of its usefulness. By focusing on individual items it was possible to overcome this. Students began to look at words with curiosity and interest, a necessary prerequisite for success as was eloquently described by Peters (1975). It did much to dispel the state of confusion with which many readers enter the task. Every effort was made to prevent the work degenerating into a series of arid and meaningless exercises and to take the drudgery of what could have become meaningless drill.

Obviously no one source of material can be regarded as comprehensive and adequate. The List referred to by Peters (1970) was found to be particularly useful (NZCER, 1963). The pupils spent most of their day in the normal classroom and the contribution of the class teacher is not to be underrated. It is felt that the continuous contribution of diagnosis to teaching helped (see Farr, 1969).

REFERENCES

AERA, Technical recommendations for achievement tests. Washington D.C.: National Education Association, 1955.

American Psychological Association, Standards for educational and psychological tests and manuals. Washington D.C.: American Psychological Association, Inc., 1966.

Anastasi, A. Psychological testing. (3rd ed.) New York: Macmillan, 1968.

Baker, F.B. Advances in item analysis. Review of Educational Research, 1977, 47(1), 151-178.

Bookbinder, G.E. Reading ages and standardized scores. Reading, 1976, 10(3), 20-23.

Clark, M.M. Young fluent readers. London: Heinemann Educational, 1976.

Closs, S.J. Ipsative vs. normative interpretation of interest test scores. Bulletin of the British Psychological Society, 1976, 29, 289-299.

Doe, B. Testing time for kids. Times Educational Supplement 19.8.77.

Eysenck, H.J. Intelligence assessment: a theoretical and experimental approach. British Journal of Educational Psychology, 1967, 37, 81-98.

Farr, R. Reading: What can be measured? Newark, Delaware: International Reading Association, 1968.

Frostig, M. Frostig developmental test of visual perception. Consulting Psychologists Press, 1963.

Godfrey Thomson Unit, University of Edinburgh, Edinburgh Reading Tests, Stage 2. London: Hodder & Stoughton, 1972.

Hieronymus, A., Lindquist, E., & France, N. Richmond test of basic skills. London: Nelson, 1974,1975.

Kirk, S.A., Mc Carthy, J., & Kirk, W. Illinois test of psycholinguistic abilities. (Rev. ed.) Urbana, Ill.: University of Illinois Press, 1968.

Moray House College of Education, Edinburgh reading tests, Stage 3, London: Hodder and Stoughton, 1972.

Neilsen, H.H., & Ringe, K. Visuo-perceptive and visuo-motor performance of children with reading disabilities. Scand. J. Psych., 1969, 10, 225. Cited by Tew (1976).

New Zealand Council for Educational Research Alphabetical Spelling List. Exeter: Wheaton, 1963.

Olson, A.V. Factor analytic studies of the Frostig developmental test of visual perception. Journal of Special Education, 1968, 2, 429.

Paraskevopoulos, J.N., & Kirk, S.A. The developmental and psychometric characteristics of the revised Illinois test of psycholinguistic abilities. Urbana, Ill.: University of Illinois Press, 1969.

Peters, M.L. Success in spelling. Cambridge: Institute of Education, 1970.

Peters, M.L. Diagnostic and remedial spelling manual. London: Macmillan, 1975.

Pumfrey, P.D. Reading: tests and assessment techniques. London: Hodder and Stoughton, 1976.

Pumfrey, P.D. Reading measurement and evaluation: some current concerns and promising developments. In J. Gilliland (Ed.), Reading: research and classroom practice. London: Ward Lock Educational for United Kingdom Reading Association, 1977.

Savage, R.D. Psychometric assessment of the individual child. Harmondsworth: Penguin Books, 1968.

Smith, P.A., & Marx, R.W. Some cautions on the use of the Frostig test. Journal of Learning Disabilities, 1972, 6, 357.

Tew, B. Some doubts about the Frostig test of visual perception. Remedial Education, 1976, 11(1), 32-35.

Topping, K.J. Correspondence. Remedial Education, 1976, 11(2), 89.

Vincent, D. Reading ages and NFER reading tests. Educational Research, 1974, 16(3), 176-180.

Vincent, D, & Cresswell, M. Reading tests in the classroom. Windsor, Berks.: National Foundation for Educational Research Publishing Company, 1976.

Vygotsky, L.S. Thought and language. Cambridge, Mass.: M.I.T. Press, 1962.

Ward, J. On the concept of criterion-referenced measurement. British Journal of Educational Psychology, 1970, 40, 314-323.

Willmott, A.S., & Fowles, D. The objective interpretation of school performance. Windsor, Berks.: National Foundation for Educational Research Publishing Company, 1974.

Woodcock, R.W. Woodcock reading mastery tests. Circle Pines, Minnesota: American Guidance Services Ltd., 1973.